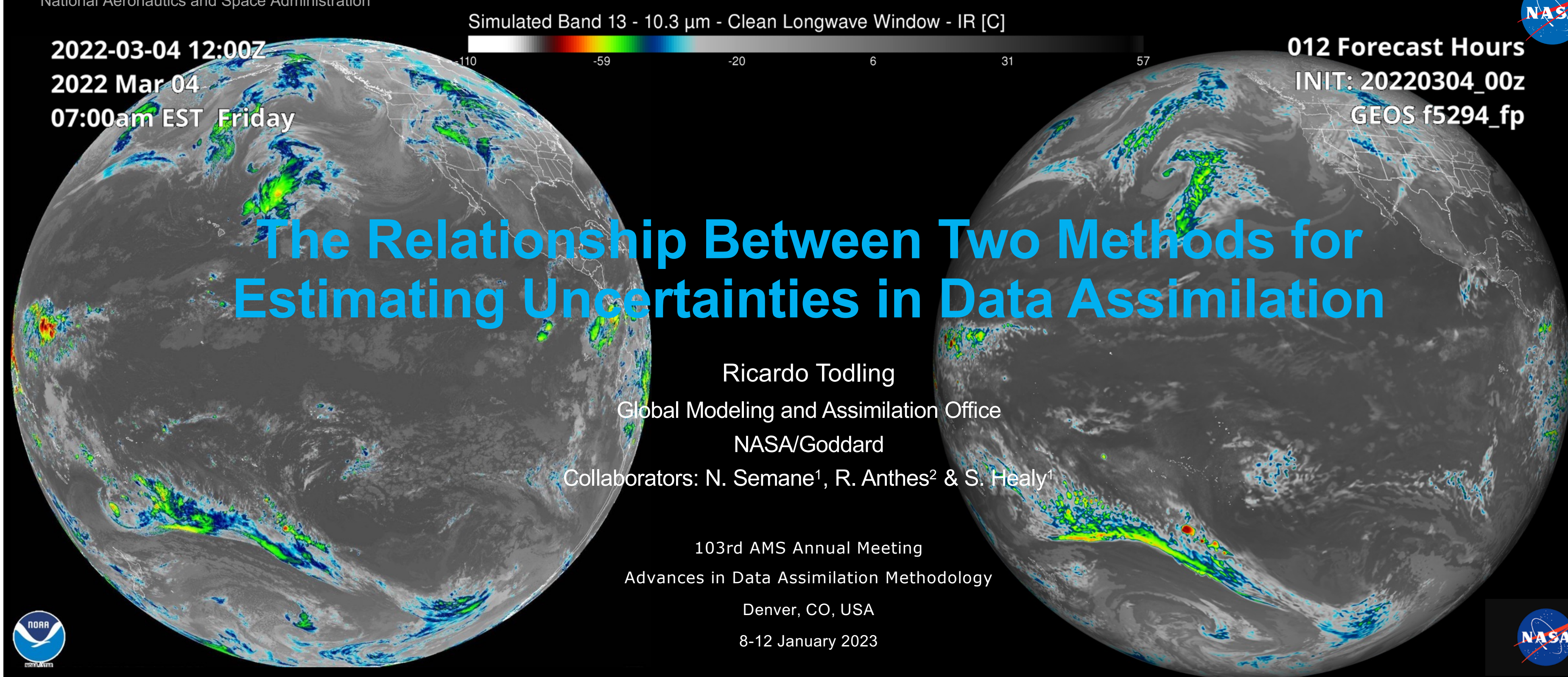
National Aeronautics and Space Administration



The Relationship Between Two Methods for Estimating Uncertainties in Data Assimilation

Ricardo Todling

Global Modeling and Assimilation Office

NASA/Goddard

Collaborators: N. Semane[1], R. Anthes[2] & S. Healy[1]

103rd AMS Annual Meeting

Advances in Data Assimilation Methodology

Denver, CO, USA

8-12 January 2023

See also R. Anthes talk in 8A.1 at 3:45PM

Revised from presentations at:

- 8th ISDA 2022
- EGU General Assembly 2022

1 ECMWF; 2 UCAR

OUTLINE

1. Motivation
2. Background
3. The relationship: 3CH & DBCP
4. Revisiting the Motivation
5. Additional Comparisons from ERA5 & GEOS
6. Closing Remarks

Motivation

In a recent work now published in J. Tech., Semane et al. (2022) compare estimates of observation uncertainty for radio occultation bending angle with the method of Desroziers et al. (2005; DBCP) and the Three-Cornered Hat (3CH) method of Gray & Allan (1974). The Fig. is a comparison.

A back of the envelop calculation during the review process (by the presenter) showed that, when things are ideal, the observation error standard deviation derived with 3CH should equal that derived with DBCP.

What is going on?

What does the mathematical analysis say?

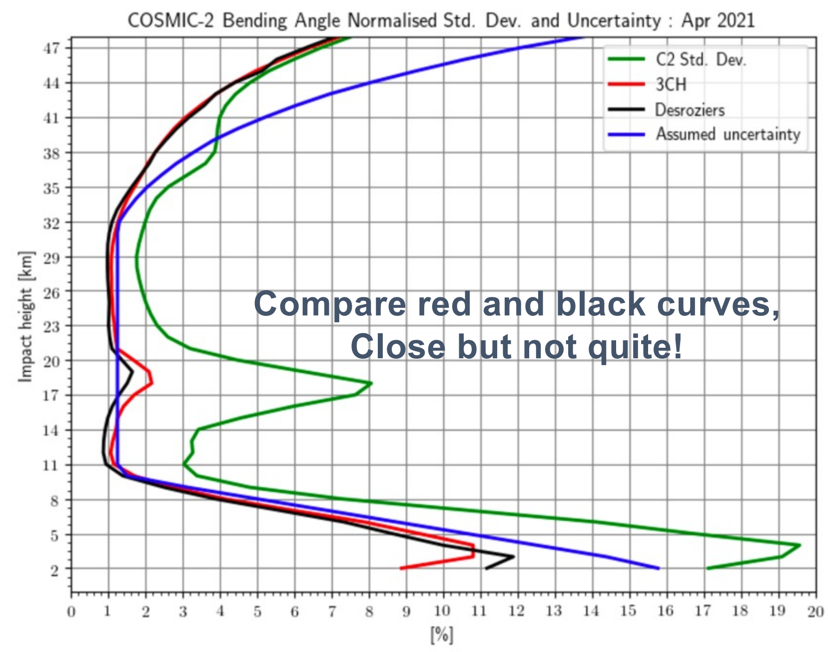


Fig. 2: Estimated COSMIC-2 bending angle random error standard deviations (uncertainties) from the Desroziers (black) and 3CH (red) methods for April 2021. The assumed ECMWF uncertainty model is shown in blue. The standard deviations of the COSMIC-2 bending angles are shown by the green profile. These estimates are for all COSMIC-2 latitudes (50°S-50°N).

Background: Cornered-Hat Methods

Atomic Timekeeping and the Statistics of Precision Signal Generators

JAMES A. BARNES

If three oscillators are used, it is possible to independently measure the three quantities σ_{12}, σ_{13}, and σ_{23}. Thus there exist three independent equations:

$$\left.\begin{aligned} \sigma_{12}^2 &= \sigma_1^2 + \sigma_2^2 \\ \sigma_{13}^2 &= \sigma_1^2 + \sigma_3^2 \\ \sigma_{23}^2 &= \sigma_2^2 + \sigma_3^2 \end{aligned}\right\}. \quad (18)$$

A METHOD FOR ESTIMATING THE FREQUENCY STABILITY OF AN INDIVIDUAL OSCILLATOR

James E. Gray and David W. Allan

Time and Frequency Division
National Bureau of Standards
Boulder

JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 103, NO. C4, PAGES 7755–7766, APRIL 15, 1998

Toward the true near-surface wind speed: Error modeling and calibration using triple collocation

Ad Stoffelen

Royal Netherlands Meteorological Institute, de Bilt, Netherlands

MARCH 2021 SJOBERG ET AL. 555

The Three-Cornered Hat Method for Estimating Error Variances of Three or More Atmospheric Datasets. Part I: Overview and Evaluation

JEREMIAH P. SJOBERG,[a] RICHARD A. ANTHES,[a] AND THERESE RIECKH[a]

[a] COSMIC Program Office, University Corporation for Atmospheric Research, Boulder, Colorado

The idea in so-called Cornered Hat Methods is to use more than one dataset of the same observable to try to estimate the uncertainty in their estimates obtained from them by taking the truth out of the way.

$$\begin{aligned} \mathbf{x} &= \mathbf{t} + \boldsymbol{\epsilon}^{\boldsymbol{x}} \\ \mathbf{y} &= \mathbf{t} + \boldsymbol{\epsilon}^{\boldsymbol{y}} \end{aligned}$$

From where it follows:

$$\begin{aligned} cov(\mathbf{t}) &= \frac{1}{4}[cov(\mathbf{x}+\mathbf{y}) - cov(\mathbf{x}-\mathbf{y})] \\ &\quad -\{E[\mathbf{t} \odot (\boldsymbol{\epsilon}^{\boldsymbol{x}} + \boldsymbol{\epsilon}^{\boldsymbol{y}})] + E(\boldsymbol{\epsilon}^{\boldsymbol{x}} \odot \boldsymbol{\epsilon}^{\boldsymbol{y}})\} \end{aligned}$$

Assuming the datasets have uncorrelated errors, an estimate of the sought uncertainties can be shown to be:

$$\begin{aligned} \hat{\mathbf{X}} &= cov(\mathbf{x}) - \frac{1}{4}[cov(\mathbf{x}+\mathbf{y}) - cov(\mathbf{x}-\mathbf{y})] \\ \hat{\mathbf{Y}} &= cov(\mathbf{y}) - \frac{1}{4}[cov(\mathbf{x}+\mathbf{y}) - cov(\mathbf{x}-\mathbf{y})] \end{aligned}$$

This is the gist of the so-called 2CH – which by neglecting the cross-term between the truth and the errors turns out to have poor accuracy (Sjoberg et al. 2021). Higher order Cornered-Hat Methods only require there be no error correlation among the chosen datasets.

Background: 3CH & DBCP

3CH Method

Given three datasets $\{\mathcal{X}, \mathcal{Y}, \mathcal{Z}\}$, the 3CH method uncertainty estimates are given by:

$$
\begin{aligned}
\hat{\mathbf{X}} &= \frac{1}{2}\{cov(\mathbf{x}-\mathbf{y}) + cov(\mathbf{x}-\mathbf{z}) - cov(\mathbf{y}-\mathbf{z})\} \\
&+ \Delta\mathbf{X} \\
\hat{\mathbf{Y}} &= \frac{1}{2}\{cov(\mathbf{y}-\mathbf{z}) + cov(\mathbf{y}-\mathbf{x}) - cov(\mathbf{z}-\mathbf{x})\} \\
&+ \Delta\mathbf{Y} \\
\hat{\mathbf{Z}} &= \frac{1}{2}\{cov(\mathbf{z}-\mathbf{x}) + cov(\mathbf{z}-\mathbf{y}) - cov(\mathbf{x}-\mathbf{y})\} \\
&+ \Delta\mathbf{Z}
\end{aligned}
$$

where $cov(\mathbf{u}, \mathbf{v}) = E[(\mathbf{u} - E(\mathbf{u}))(\mathbf{v} - E(\mathbf{v}))^T]$, and with

$$
\begin{aligned}
\Delta\mathbf{X} &= E(\boldsymbol{\epsilon}^{x} \odot \boldsymbol{\epsilon}^{y}) + E(\boldsymbol{\epsilon}^{x} \odot \boldsymbol{\epsilon}^{z}) - E(\boldsymbol{\epsilon}^{y} \odot \boldsymbol{\epsilon}^{z}) \\
\Delta\mathbf{Y} &= E(\boldsymbol{\epsilon}^{y} \odot \boldsymbol{\epsilon}^{z}) + E(\boldsymbol{\epsilon}^{y} \odot \boldsymbol{\epsilon}^{x}) - E(\boldsymbol{\epsilon}^{z} \odot \boldsymbol{\epsilon}^{x}) \\
\Delta\mathbf{Z} &= E(\boldsymbol{\epsilon}^{z} \odot \boldsymbol{\epsilon}^{x}) + E(\boldsymbol{\epsilon}^{z} \odot \boldsymbol{\epsilon}^{y}) - E(\boldsymbol{\epsilon}^{x} \odot \boldsymbol{\epsilon}^{y})
\end{aligned}
$$

being the *unaccessible* random terms.

Practical use of 3CH looks for three datasets with independent errors, so the "delta" terms can be safely disregarded.

DBCP Method

Desroziers et al. (2005; DBCP)

$$
\begin{aligned}
\hat{\mathbf{R}} &\equiv \frac{1}{2}[cov(\mathbf{o}-\mathbf{a}, \mathbf{o}-\mathbf{b}) + cov(\mathbf{o}-\mathbf{b}, \mathbf{o}-\mathbf{a})] \\
\hat{\mathbf{B}} &\equiv \frac{1}{2}[cov(\mathbf{a}-\mathbf{b}, \mathbf{o}-\mathbf{b}) + cov(\mathbf{o}-\mathbf{b}, \mathbf{a}-\mathbf{b})] \\
\hat{\mathbf{A}} &\equiv \frac{1}{2}[cov(\mathbf{a}-\mathbf{b}, \mathbf{o}-\mathbf{a}) + cov(\mathbf{o}-\mathbf{a}, \mathbf{a}-\mathbf{b})]
\end{aligned}
$$

where, for simplicity, $\mathbf{b} = \mathbf{H}\mathbf{x}^b$ and $\mathbf{a} = \mathbf{H}\mathbf{x}^a$.

Under $\tilde{\mathbf{B}} + \tilde{\mathbf{R}} \overset{icc}{=} \mathbf{B} + \mathbf{R}$	Full Optimality
$\hat{\mathbf{R}} \overset{icc}{=} \tilde{\mathbf{R}}$	$\hat{\mathbf{R}} \overset{opt}{=} \mathbf{R}$
$\hat{\mathbf{B}} \overset{icc}{=} \tilde{\mathbf{B}}$	$\hat{\mathbf{B}} \overset{opt}{=} \mathbf{B}$
$\hat{\mathbf{A}} \overset{icc}{=} \tilde{\mathbf{A}}$	$\hat{\mathbf{A}} \overset{opt}{=} \mathbf{A}$

where $\tilde{\mathbf{A}} = (\mathbf{I} - \tilde{\mathbf{B}}\tilde{\Gamma}^{-1})\tilde{\mathbf{B}}$, and $\tilde{\Gamma} = \tilde{\mathbf{B}} + \tilde{\mathbf{R}}$. $\tilde{\mathbf{B}}$ and $\tilde{\mathbf{R}}$ are the *prescribed* background and observation error covariances; $\tilde{\mathbf{A}}$ is the *perceived* analysis error covariance; and where $\mathbf{R}$, $\mathbf{B}$ and $\mathbf{A}$ are *true* error covariances.

Comparison: 3CH vs DBCP

3CH Method

Have 3CH choose $\{\mathcal{O},\mathcal{B},\mathcal{A}\}$ for its $\{\mathcal{X},\mathcal{Y},\mathcal{Z}\}$ datasets:

$$
\begin{aligned}
\hat{\mathbf{X}} &= \frac{1}{2}\{cov(\mathbf{o}-\mathbf{b}) + cov(\mathbf{o}-\mathbf{a}) - cov(\mathbf{b}-\mathbf{a})\} \\
&+ \Delta\mathbf{X} \\
\hat{\mathbf{Y}} &= \frac{1}{2}\{cov(\mathbf{b}-\mathbf{a}) + cov(\mathbf{b}-\mathbf{o}) - cov(\mathbf{a}-\mathbf{o})\} \\
&+ \Delta\mathbf{Y} \\
\hat{\mathbf{Z}} &= \frac{1}{2}\{cov(\mathbf{a}-\mathbf{o}) + cov(\mathbf{a}-\mathbf{b}) - cov(\mathbf{o}-\mathbf{b})\} \\
&+ \Delta\mathbf{Z}
\end{aligned}
$$

with

$$
\begin{aligned}
\Delta\mathbf{X} &= E(\epsilon^o \odot \epsilon^b) + E[\epsilon^a \odot (\epsilon^o - \epsilon^b)] \\
\Delta\mathbf{Y} &= E(\epsilon^o \odot \epsilon^b) - E[\epsilon^a \odot (\epsilon^o - \epsilon^b)] \\
\Delta\mathbf{Z} &= E[\epsilon^a \odot (\epsilon^o + \epsilon^b)] - E(\epsilon^o \odot \epsilon^b)
\end{aligned}
$$

being the *unaccessible* random terms.

At first sight, this would seem to be an odd choice of corners since $\mathbf{a}$ is correlated with $\mathbf{o}$ and $\mathbf{b}$. In this case, what does 3CH get?

x ⇔ o
y ⇔ b
z ⇔ a

DBCP Method

Desroziers et al. (2005; DBCP)

$$
\begin{aligned}
\hat{\mathbf{R}} &\equiv \frac{1}{2}[cov(\mathbf{o}-\mathbf{a},\mathbf{o}-\mathbf{b}) + cov(\mathbf{o}-\mathbf{b},\mathbf{o}-\mathbf{a})] \\
\hat{\mathbf{B}} &\equiv \frac{1}{2}[cov(\mathbf{a}-\mathbf{b},\mathbf{o}-\mathbf{b}) + cov(\mathbf{o}-\mathbf{b},\mathbf{a}-\mathbf{b})] \\
\hat{\mathbf{A}} &\equiv \frac{1}{2}[cov(\mathbf{a}-\mathbf{b},\mathbf{o}-\mathbf{a}) + cov(\mathbf{o}-\mathbf{a},\mathbf{a}-\mathbf{b})]
\end{aligned}
$$

where, for simplicity, $\mathbf{b} = \mathbf{H}\mathbf{x}^b$ and $\mathbf{a} = \mathbf{H}\mathbf{x}^a$.

Under $\tilde{\mathbf{B}} + \tilde{\mathbf{R}} \overset{icc}{=} \mathbf{B} + \mathbf{R}$	Full Optimality
$\hat{\mathbf{R}} \overset{icc}{=} \tilde{\mathbf{R}}$	$\hat{\mathbf{R}} \overset{opt}{=} \mathbf{R}$
$\hat{\mathbf{B}} \overset{icc}{=} \tilde{\mathbf{B}}$	$\hat{\mathbf{B}} \overset{opt}{=} \mathbf{B}$
$\hat{\mathbf{A}} \overset{icc}{=} \tilde{\mathbf{A}}$	$\hat{\mathbf{A}} \overset{opt}{=} \mathbf{A}$

where $\tilde{\mathbf{A}} = (\mathbf{I} - \tilde{\mathbf{B}}\tilde{\Gamma}^{-1})\tilde{\mathbf{B}}$, and $\tilde{\Gamma} = \tilde{\mathbf{B}} + \tilde{\mathbf{R}}$. $\tilde{\mathbf{B}}$ and $\tilde{\mathbf{R}}$ are the *prescribed* background and observation error covariances; $\tilde{\mathbf{A}}$ is the *perceived* analysis error covariance; and where $\mathbf{R}$, $\mathbf{B}$ and $\mathbf{A}$ are *true* error covariances.

Comparison: 3CH cross-terms

Given that A is correlated with O & B errors, why should {*O,B,A*} be a viable choice of corners?

Answer: lucky when it comes to the first two corners:

- O & B errors are (assumed) uncorrelated.
- Analysis errors are orthogonal to O-B residuals.

However, this is not so for the third corner:

- Random error adds up to twice the analysis error covariance.

$$
\begin{aligned}
\Delta\mathrm{X} &= E(\epsilon^o \odot \epsilon^b) + E[\epsilon^a \odot (\epsilon^o - \epsilon^b)] \\
\Delta\mathrm{Y} &= E(\epsilon^o \odot \epsilon^b) - E[\epsilon^a \odot (\epsilon^o - \epsilon^b)] \\
\Delta\mathrm{Z} &= E[\epsilon^a \odot (\epsilon^o + \epsilon^b)] - E(\epsilon^o \odot \epsilon^b)
\end{aligned}
$$

- Uncorrelated O & B errors: $E(\epsilon^o \odot \epsilon^b) = 0$
- Orthogonality: $E[\epsilon^a \odot (\epsilon^o - \epsilon^b)] = 0$

Therefore,

$$
\begin{aligned}
\Delta\mathrm{X} &= 0 \\
\Delta\mathrm{Y} &= 0 \\
\Delta\mathrm{Z} &= 2E[\epsilon^a \odot \epsilon^b)] \\
&\overset{icc}{=} 2\tilde{\mathrm{A}} \leftarrow \text{innovation covariance consistency} \\
&\overset{opt}{=} 2\mathrm{A} \leftarrow \text{optimal}
\end{aligned}
$$

The Relationship: 3CH & DBCP

With the association: $\{\mathcal{X},\mathcal{Y},\mathcal{Z}\} \rightarrow \{\mathcal{O},\mathcal{B},\mathcal{A}\}$:

Full 3CH

Suboptimal case:

$$\begin{aligned}\hat{\mathbf{X}} &= \tilde{\Gamma} - \mathbf{B}\\ \hat{\mathbf{Y}} &= \mathbf{B}\\ \hat{\mathbf{Z}} &= (\mathbf{I}-\tilde{\mathbf{K}})\mathbf{B}(\mathbf{I}-\tilde{\mathbf{K}})^T + \tilde{\mathbf{K}}\mathbf{R}\tilde{\mathbf{K}}^T\end{aligned}$$

for $\tilde{\mathbf{K}} = \tilde{\mathbf{B}}\tilde{\Gamma}^{-1}$, i.e., $\hat{\mathbf{Z}}$ arrives at Joseph's formula for the *actual* analysis error covariance (filter performance).

Under Optimality: 3CH = DBCP.

Practical 3CH: neglect of cross–terms

Under Innovation Covariance Consistency:

$$\begin{aligned}\hat{\mathbf{X}} &\stackrel{icc}{=} \tilde{\mathbf{R}}\\ \hat{\mathbf{Y}} &\stackrel{icc}{=} \tilde{\mathbf{B}}\\ \hat{\mathbf{Z}} &\stackrel{icc}{=} -\tilde{\mathbf{A}}\end{aligned}$$

Under Optimality:

$$\begin{aligned}\hat{\mathbf{X}} &\stackrel{opt}{=} \mathbf{R}\\ \hat{\mathbf{Y}} &\stackrel{opt}{=} \mathbf{B}\\ \hat{\mathbf{Z}} &\stackrel{opt}{=} -\mathbf{A}\end{aligned}$$

Why do these results follow?

In addition to the fact that observation and background errors are assumed uncorrelated …

When it comes to the first two corners of 3CH:

- Random errors cancel out due to the orthogonality between the analysis error and the innovation vector.

When it comes to the third corner of 3CH:

- Random error add up to twice the analysis error covariance.

Revisiting the Motivation

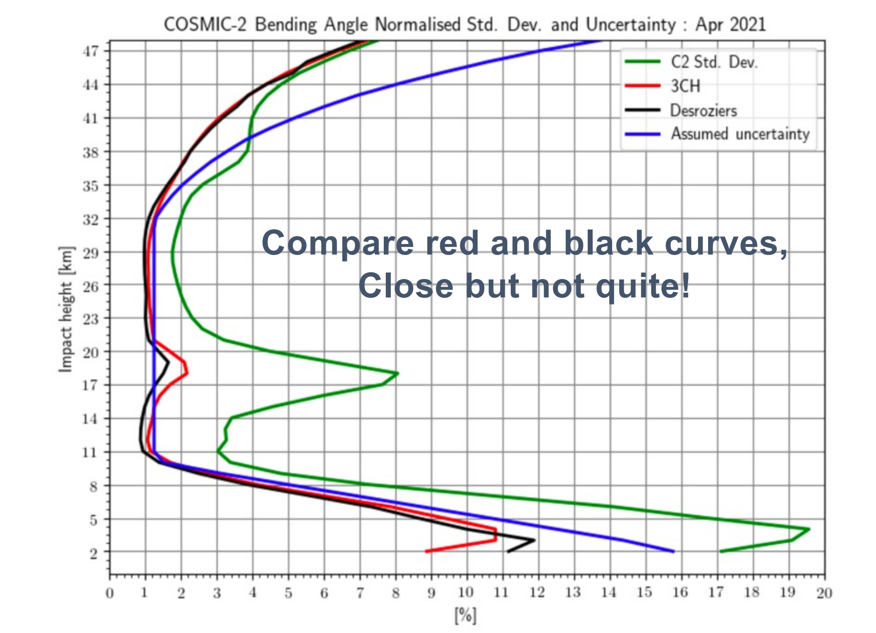
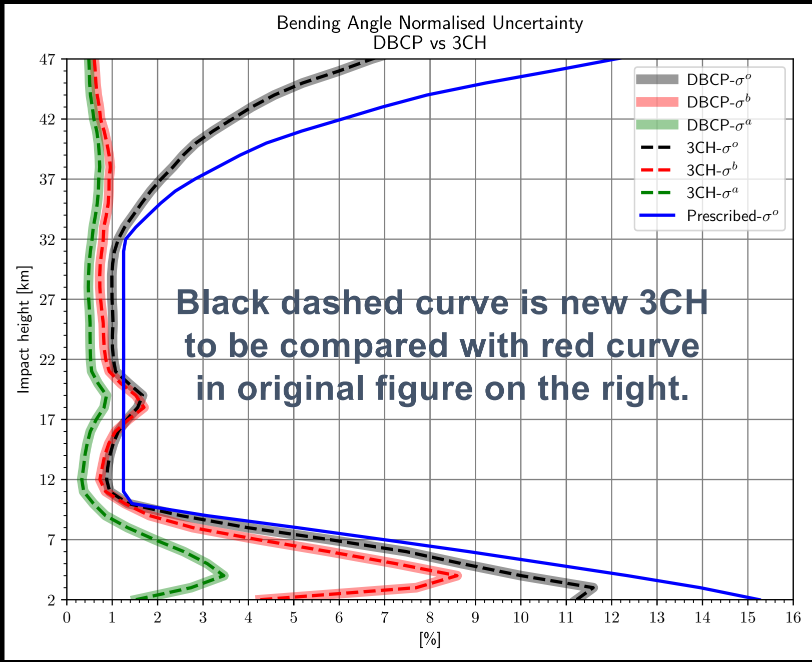


Fig. 2: Estimated COSMIC-2 bending angle random error standard deviations (uncertainties) from the Desroziers (black) and 3CH (red) methods for April 2021. The assumed ECMWF uncertainty model is shown in blue. The standard deviations of the COSMIC-2 bending angles are shown by the green profile. These estimates are for all COSMIC-2 latitudes (50°S-50°N).

Note: only black and blue curves should be compared across plots; other curves are for different quantities.

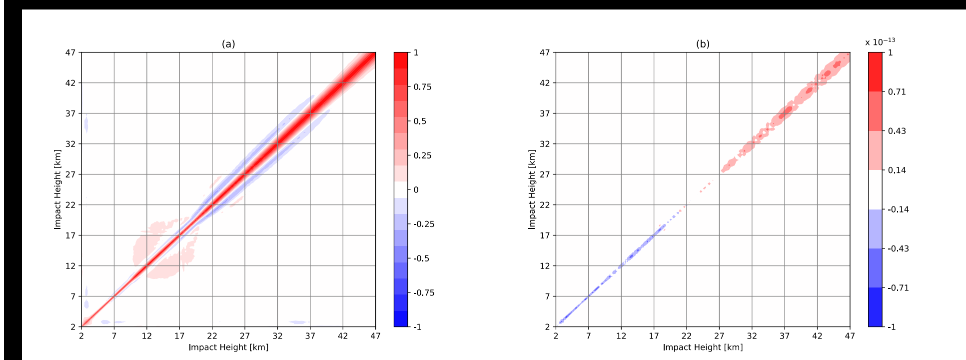


- Re-calculating 3CH using fully consistent corners as compared to DBCP residual vectors from ERA5.
- Diffs between 3CH & DBCP in Semane et al. (2022) are simply due to sampling strategy.

Additional Comps: Observation Uncertainty Correlations

Results from ERA5 residuals for April 2021

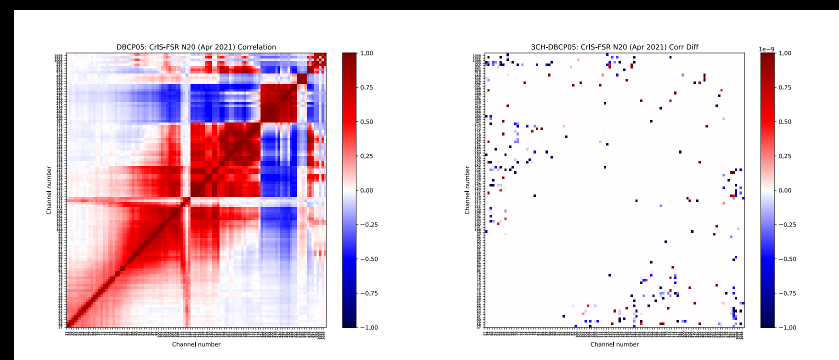
Vertical Correlations: RO Bending Angle



DBCP **Diff(DBCP,3CH) ~ 10^{-13}**

Results from NASA-GEOS residuals for April 2021

Interchannel Correlations for CrIS-FSR NOAA-20



DBCP **Diff(DBCP,3CH) ~ 10^{-9}**

The methods produce nearly (round-off) identical results,
When the datasets used in 3CH are consistent with the residuals of DBCP.

Closing Remarks

- This work shows that a particular choice of datasets makes the Three-Cornered Hat Method (3CH) results to be identical to those of Desroziers et al. (2005; DBCP).
- The full version of 3CH requires knowledge of random errors not known in practice.
- When 3CH neglects the unknow terms, two of its corners recover identical results as DBCP for the observation and background error covariances; a lucky coincidence due to:
  - their errors being uncorrelated; and
  - the analysis errors being orthogonal to the innovation vector.
- Surprisingly, when the random terms are neglected, the third corner of 3CH recovers the *negative* of the analysis error covariance, in contrast to DBCP.